

Copernicus FICE 2024

Training on


In situ Ocean Colour Above-Water Radiometry towards Satellite Validation


Quantification of uncertainties in satellite data products


Giuseppe Zibordi


giuseppe.zibordi@eoscience.eu



PROGRAMME OF THE EUROPEAN UNION  Copernicus
Europe's eyes on Earth

IMPLEMENTED BY  EUMETSAT

FRM4SOC Phase-2  fiducial reference measurements for satellite ocean colour

 CNR ISMAR
ISTITUTO DI SCIENZE MARINE

6-17 May 2024
Venice, Italy

Matchups construction

Matchups are constructed applying satellite extraction protocols, which may be different and naturally lead to appreciable differences in matchups results.

As an example, matchups can be confidently constructed using the median of the 3×3 satellite pixels centred at various measurement sites and applying criteria to maximize the *comparability* of *satellite* and *in situ* data.

Matchups can be retained for successive analysis when:

- i. the time difference Δt between *in situ* measurement and satellite overpass is less than ± 2 hr (only holding the *in situ* data closest in time to the satellite overpass);
- ii. none of the 3×3 pixels is affected by the standard processing flags;
- iii. the coefficient of variation (*i.e.*, the ratio of standard deviation to mean) of $L_{WN}^{OLCI-A}(\lambda)$ is lower than 20% at 560 nm for the nine pixels (the 560 nm centre-wavelength is expected to exhibit a lower dependence on optically significant constituents and surface perturbations with respect to the other centre-wavelengths in the visible portion of the spectrum);
- iv. the viewing angle is lower than 60° ;
- v. the sun zenith angle is lower than 70° ; and
- vi. the aerosol optical depth τ_a determined at a near-infrared centre-wavelength (*i.e.*, 865 nm) is lower than 0.5 to avoid data affected by cloud perturbations.

Matchups statistics

Satellite data products can be evaluated through statistical indices for the N matchups of satellite (*SAT*) and *in situ* (*PRS*) data $[(\mathfrak{I}_1^{SAT}, \mathfrak{I}_2^{SAT}, \dots, \mathfrak{I}_N^{SAT}), (\mathfrak{I}_1^{PRS}, \mathfrak{I}_2^{PRS}, \dots, \mathfrak{I}_N^{PRS})]$ where \mathfrak{I} is the compared quantity (*i.e.*, $L_{WN}(\lambda)$), and the subscripts 1, 2 ..., N indicate the matchup index.

Statistical indices that can be considered are:

- i. the median of differences Δ_m and the median of absolute (unsigned) differences $|\Delta|_m$;
- ii. the median of relative differences ψ_m and the median of absolute (unsigned) relative differences $|\psi|_m$, both determined with respect to the *in situ* reference data;
- iii. the root mean square of differences *rmsd*; and the determination coefficient r^2 from data regression.

The indices $|\Delta|_m$ and $|\psi|_m$ provide hints on the dispersion of data, conversely Δ_m and ψ_m provide information on biases. The quantities $|\psi|_m$ and ψ_m are expressed in percent and provide an immediate view on the comparison. On the contrary, $|\Delta|_m$, and *rmsd* are in physical units (*e.g.*, $mW\ cm^{-2}\ \mu m^{-1}\ sr^{-1}$) and complement the comparison with statistical indices strictly related to the values and range of the assessed $L_{WN}(\lambda)$.

When considering that the $L_{WN}(\lambda)$ values may not exhibit normal distribution, the use of the *median* with respect to the *mean* allows to better determine the *centrality* of the comparison results.

The direct method: general

The comparison of satellite versus *in situ* data provides only general indications on the accuracy affecting data products: none of the statistical quantities formerly associated to the comparison is a direct expression of the uncertainties.

Accuracy requirements for satellite radiometric products are commonly summarized by a spectrally and water-type independent *5% uncertainty*. This generic *5% uncertainty* requirement, however, should only apply to oligotrophic/mesotrophic waters in the blue-green spectral region. To assess the fulfilment of such a requirement, radiometric data products can be evaluated applying the so-called consistency principle: independent measurements of the same quantity should agree within their individual uncertainties (Immler *et al.* 2010).



The direct method: application

By:

- i. applying the consistency principle to satellite $L_{WN,i}^{SAT}(\lambda)$ and *in situ* $L_{WN,i}^{PRS}(\lambda)$ matchups with i indicating the matchup index;
- ii. choosing a coverage factor $k = 1$ (indicating 100%); and
- iii. assuming negligible correlations between uncertainties, the following relationship should be statistically satisfied:

$$\left[L_{WN,i}^{SAT}(\lambda) - L_{WN,i}^{PRS}(\lambda) \right]^2 < k \times \left[u_{SAT,i}^2(\lambda) + u_{PRS,i}^2(\lambda) + v_{SAT,i}^2(\lambda) + v_{PRS,i}^2(\lambda) \right]$$

where $u_{SAT,i}(\lambda)$ indicates the expected uncertainty of satellite data, $u_{PRS,i}(\lambda)$ the quantified uncertainty for the *in situ* data and, $v_{SAT,i}(\lambda)$ and $v_{PRS,i}(\lambda)$ the spatio-temporal variabilities affecting satellite and *in situ* data, respectively.

Assuming that most of the major contributions to radiance differences are accounted for

$$\varepsilon_i(\lambda) = \frac{L_{WN,i}^{SAT}(\lambda) - L_{WN,i}^{PRS}(\lambda)}{\sqrt{u_{SAT,i}^2(\lambda) + u_{PRS,i}^2(\lambda) + v_{SAT,i}^2(\lambda) + v_{PRS,i}^2(\lambda)}}$$

Interpretation of comparison results

$$\varepsilon_i(\lambda) = \frac{L_{WN,i}^{SAT}(\lambda) - L_{WN,i}^{PRS}(\lambda)}{\sqrt{u_{SAT,i}^2(\lambda) + u_{PRS,i}^2(\lambda) + v_{SAT,i}^2(\lambda) + v_{PRS,i}^2(\lambda)}}$$

In agreement with Hunt *et al.* (2020) and following Zibordi *et al.* (2022), if $u_{SAT,i}(\lambda)$, $u_{PRS,i}(\lambda)$, $v_{SAT,i}(\lambda)$, $v_{PRS,i}(\lambda)$ well describe the variance of the differences $L_{WN,i}^{SAT}(\lambda) - L_{WN,i}^{PRS}(\lambda)$, the probability distribution of the $\varepsilon_i(\lambda)$ values would be standard normal and consequently centred at 0 with standard deviation $\sigma(\varepsilon_i(\lambda))$ equal to 1.

In the case the uncertainty values are underestimated, the distribution of $\varepsilon_i(\lambda)$ would exhibit a standard deviation $\sigma(\varepsilon_i) > 1$. On the contrary, if the uncertainty values are overestimated, the distribution of $\varepsilon_i(\lambda)$ would lead to $\sigma(\varepsilon_i) < 1$ (*see the equation for ε_i*).

Impact of Δ_t on PRS uncertainties

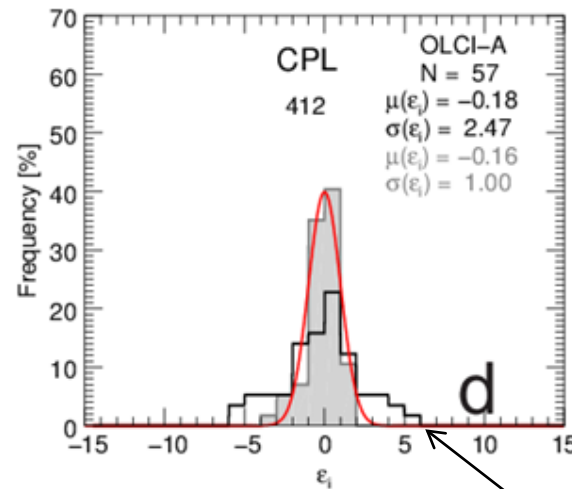
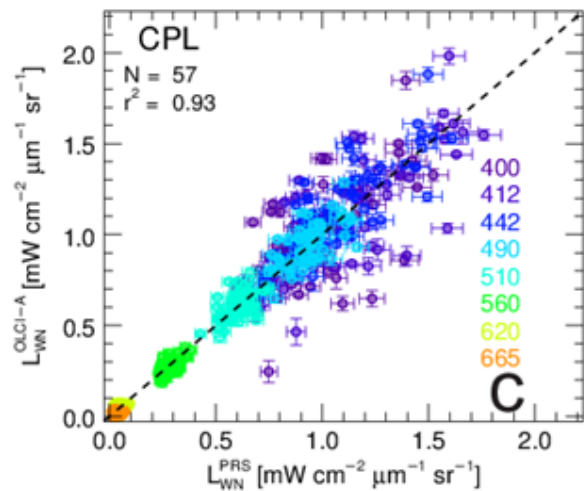
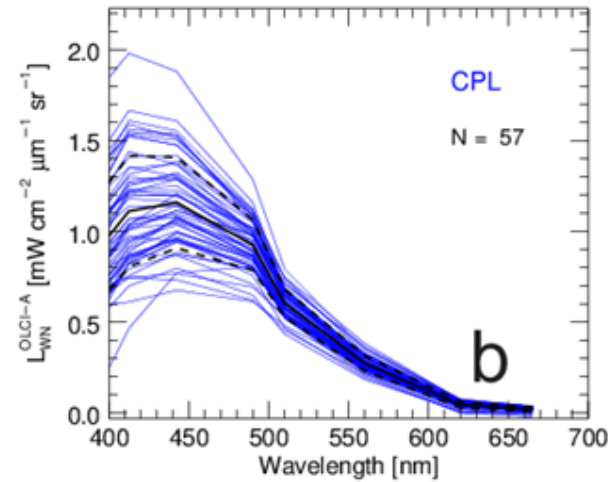
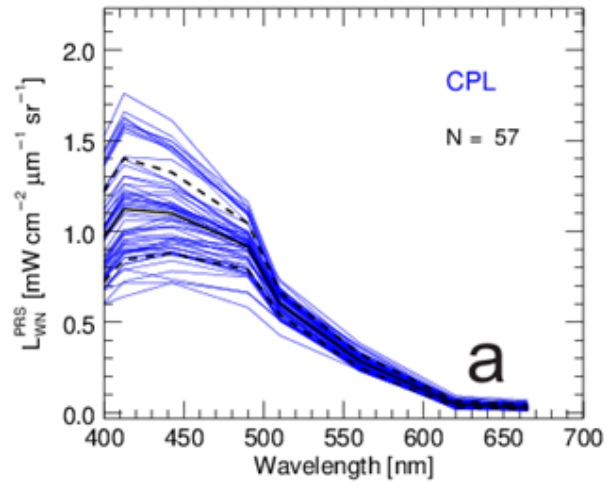
Site	Δ_t [hours]	400	412	443	490	510	560	620	667	N
CPL	1	3.5	3.3	3.0	2.8	3.0	4.1	11.3	14.4	366
	2	3.9	3.7	3.2	3.1	3.2	4.4	12.1	15.1	188
AAOT	1	3.9	3.7	3.3	3.1	3.2	4.1	9.9	12.5	483
	2	6.3	5.6	4.9	4.7	4.7	5.6	7.9	9.0	260

Median relative uncertainty $u_{mu}(L_{WN}^{Chla})/L_{WN}^{Chla}$ (in %) of AERONET-OC situ data contributing to matchups constructed with diverse Δ_t , for $\tau_a(412) \leq 0.2$ and $\theta_0 \leq 45^\circ$.

Pair of triplets within the interval Δ_t are used to estimate the spatio-temporal uncertainties characterizing the in situ data.

N is the number of pairs of triplets used to quantify the dependence on Δ_t .

Application



- (a) AERONET-OC $L_{WN}^{PRS}(\lambda)$
- (b) OLCI-A $L_{WN}^{OLCI-A}(\lambda)$
- (c) matchup spectra, scatter plot of $L_{WN}^{OLCI-A}(\lambda)$ versus $L_{WN}^{PRS}(\lambda)$ matchup data, and
- (d) distributions of the uncertainty-normalized difference ε_i at the 412 nm imposing $u_{OLCI-A, i}(\lambda) = 0.05 \cdot L_{WN, i}^{OLCI-A}(\lambda)$ (black line) and alternatively imposing a constant value for $u_{OLCI-A, i}(\lambda)$ so that $\sigma(\varepsilon_i(\lambda)) = 1$ (grey line and shaded background) compared to an ideal normal distribution (red line).

N indicates the number of matchups, r^2 the determination coefficient and, $\mu(\varepsilon_i)$ and $\sigma(\varepsilon_i)$ the mean and standard deviation of the ε_i values, respectively.

$\sigma(\varepsilon_i) > 1$. Thus the 5% uncertainty is underestimated!

$$\varepsilon_i(\lambda) = \frac{L_{WN,i}^{SAT}(\lambda) - L_{WN,i}^{PRS}(\lambda)}{\sqrt{u_{SAT,i}^2(\lambda) + u_{PRS,i}^2(\lambda) + v_{SAT,i}^2(\lambda) + v_{PRS,i}^2(\lambda)}}$$

λ [nm]	400	412	442	490	560	665
$\mu(L_{WN,i}^{PRS}) \pm \sigma(L_{WN,i}^{PRS})$ [mW cm ⁻² μm ⁻¹ sr ⁻¹]	0.972±0.248	1.123±0.279	1.100±0.224	0.916±0.128	0.285±0.038	0.030±0.014
$\mu(u_{PRS,i}/L_{WN,i}^{PRS}) \pm \sigma(u_{PRS,i}/L_{WN,i}^{PRS})$	0.050±0.002	0.049±0.002	0.049±0.002	0.051±0.002	0.073±0.004	0.404±0.163
$\mu(v_{SAT,i}/L_{WN,i}^{SAT}) \pm \sigma(v_{SAT,i}/L_{WN,i}^{SAT})$	0.035±0.036	0.031±0.024	0.025±0.016	0.020±0.012	0.039±0.021	0.469±0.988
$ \psi $	0.144	0.130	0.088	0.070	0.104	0.434
ψ	-0.002	-0.007	+0.026	+0.005	-0.037	-0.227
$\mu(\varepsilon_i(u_{SAT,i} = 0.05 \cdot L_{WN,i}^{SAT})) \pm \sigma(\varepsilon_i)$	-0.06±2.90	-0.18±2.47	+0.66±1.67	+0.12±1.31	-0.41±1.32	-0.88±1.47
$\mu(\varepsilon_i(\sigma(\varepsilon_i) = 1)) \pm \sigma(\varepsilon_i(\sigma(\varepsilon_i) = 1))$	-0.15±1.00	-0.16±1.00	+0.36±1.00	+0.07±1.00	-0.35±1.00	-0.66±1.00
$\mu(u_{SAT,i}(\varepsilon_i(\sigma(\varepsilon_i) = 1))/L_{WN,i}^{SAT})$	0.302	0.203	0.107	0.080	0.100	0.920

Fundamental sentences



‘Good (practically useful) data do not collect themselves. Neither do they magically appear on one’s desk, ready for analysis and lending insight into how to improve processes’ (Vardemann and Jobe 2016)

‘... adequately sampled, carefully calibrated, quality controlled, and archived data for key elements of the climate system will be useful indefinitely’ (Wunsch et al. 2013)

‘... a measurement of any kind is incomplete unless accompanied with an estimate of the uncertainty associated with that measurement ‘ (Palmer and Grant, 2010)

‘... we should do the radiometry correctly, or not do it at all’ (Richard Beck, 2022)

Quantification of uncertainties in satellite derived radiometric data-products

Outline

- Common practices
 - Matchup construction (basic criteria)
 - Matchup analysis (basic stats)
- The direct method
- In situ uncertainties (radiometric and spatial/temporal, parameterization)
- An application to satellite data (VIIRS, OLCI)
- General considerations

